

## Introducing the Israeli Supreme Court Database

KEREN WEINSHALL *The Hebrew University of Jerusalem*

LEE EPSTEIN *Washington University in St. Louis*

ANDY WORMS

*Advancing knowledge and driving discovery often require data infrastructure. To that end, we built the Israeli Supreme Court Database (ISCD), which encodes information from all final decisions of Israeli Supreme Court in cases opened between 2010 and 2018 (16,109 cases and 48,634 opinions). Guiding our work were the five defining characteristics of high-quality infrastructure, such that the Database is accessible, reliable and reproducible, sustainable, foundational, and capable of addressing real-world problems. In what follows we elaborate on these criteria, offer examples, and, more generally, introduce the ISCD to the research, policy, and legal communities.*

Word Count: 1093

**T**he past few years have witnessed dramatic growth in the empirical analysis of apex courts—the Israeli Supreme Court (ISC) not excepted. Since the early 2000s, hundreds of quantitative studies of the ISC have appeared in scholarly journals and newspapers.<sup>1</sup> Although the studies focus on different aspects of the Court’s work, they share a reliance on one-off (or otherwise limited) hand-coded datasets designed to assess particular hypotheses. The “one-off”

---

Keren Weinshall (keren.weinshall@mail.huji.ac.il) is the Katia & Hans Guth-Dreyfus Chair in Conflict Resolution and Law at The Hebrew University in Jerusalem. Lee Epstein (epstein@wustl.edu) is the Ethan A.H. Shepley Distinguished University Professor at Washington University in St. Louis. The Israeli Supreme Court Database has been generously supported by the I-CORE program of the Planning and Budgeting Committee and the Israel Science Foundation (grant no. 1821/12). Epstein thanks the the Guggenheim Foundation, the U.S. National Science Foundation, and Washington University School of Law for supporting her work on judicial behavior.

**Introduction prepared for the Comparative Supreme Court Decision Making Workshop**

<sup>1</sup>For literature reviews, see Esienberg et al. (2011); Teichman and Zamir (2014); Weinshall-Margel (2016)

approach has its benefits of course. But that it also has substantial costs, including massive duplication of effort, inefficiencies, dated information and measures, and conflicting results. These costs, we believe, ultimately impede the drive to discovery.

For these reasons, we built the Israeli Supreme Court Dataset (ISCD): data infrastructure designed to advance knowledge and accelerate innovation by encoding information from all final decisions of the ISC in primary cases opened between 2010 and 2018 (16,109 cases and 48,634 opinions). More specifically, the ISCD consists of 61 variables (columns) for each case and 74 variables for each justice's opinion in the population of cases decided by panels of three to nine justices. The variables capture information on the parties, litigants and legal representation, the origin and history of appealed cases, proceedings and hearings in the ISC, case outcomes, and the opinions and background characteristics of the individual justices.

In building the dataset, we were guided by the five defining characteristics of high-quality infrastructure.<sup>2</sup> First, members of the community should be able access it with no barriers to entry or use. To this end, not only is the ISCD one of the very few non-U.S. based databases that is freely and publicly available; its website also houses an analysis tool that allows users to access particular variables without having to download the dataset.

Second, high-quality infrastructure should be reproducible and the encoded data, reliable. Reproducibility means that the developers and the users must understand how to reproduce the data housed in the infrastructure. Reliability is related: It is the extent to which it is possible to replicate encoded data, producing the same value using the same standard for the same subject at the same time regardless of who or what is doing the replicating. Through a combination of automated and hand coding (with reliability assessed) we sought to meet both criteria.<sup>3</sup>

Standing the test of time is the third characteristic of high-quality infrastructure; and we have developed several strategies to meet it. Chiefly, to the extent possible we eliminated humans from the

---

<sup>2</sup>Adapted from Epstein (2019); Epstein et al. (2019); Epstein and Martin (2014).

<sup>3</sup>Because we provide more details in the body of the paper, suffice it to note here: (1) the Database's documentation provides the precise definition and coding method for each variable; and (second, that (2) various computerized and human reliability tests show over 90% accuracy/agreement.

---

data-generation process, as we just suggested. We also sought to repel (irrational) data exuberance on the theory that intricate and detailed coding schemes are the surest way to create a product that will die a slow death, not to mention unreliable data.<sup>4</sup>

Which brings us to the fourth and related characteristic: Data infrastructure should serve as foundation on which researchers can build by adding content, backdating, updating, or otherwise adapting it to their own needs; the infrastructure need not—and more to the point, should not—be the be-all, end-all. To this end we sought to build a product that, yes, will be useful in its own right or even for purposes of comparison with other apex courts<sup>5</sup> but that also can be adapted to future or even present needs.

Last but not least, by definition data infrastructure should promote innovation, inventions, and insights.<sup>6</sup> Although no product can guarantee these ends, infrastructure aimed at solving or developing implications for real-world problems increases the odds of success. By providing information of interest to policymakers, journalists, and citizens seeking to make evidence-based assessments of the ISC and its work, we believe the Dataset meets this criterion; and, in what follows, offer supporting examples—some of which shore up surprising trends in ISC decisions since 2010.

We now turn to elaborating on these criteria, providing illustrations, and, more generally, introducing the ISCD to the research, policy, and legal communities.

## REFERENCES

Benoit, Kenneth , Drew Conway, Benjamin E. Lauderdale, Michael Laver, and Slava Mikhaylov (2016). Crowd-sourced text analysis: Reproducible and agile production of political data. *American Political Science Review* 110, 278–295.

---

<sup>4</sup>See Benoit et al. (2016).

<sup>5</sup>In creating case-based datasets of high courts, many scholars have adapted the U.S. Supreme Court Database’s (<http://supremecourtdatabase.org>) protocols and coding scheme (see Epstein et al. 2019). The ISCD follows suit, as we describe later in the text.

<sup>6</sup>See, e.g., Epstein and King (2002); Epstein and Martin (2014).

- Epstein, Lee (2019). Forward movement in the development of data infrastructure for the comparative analysis of law and legal institutions. In D. Kapiszewski and M. C. Ingram (Eds.), *Concepts, Data, and Methods in Comparative Law and Politics*. Cambridge University Press, forthcoming.
- Epstein, Lee , James L. Gibson, and Andrew D. Martin (2019). Using databases to study constitutional law. In D. Law (Ed.), *Handbook of Research Methods in Constitutional Law*. Edward Elgar, forthcoming.
- Epstein, Lee and Gary King (2002). The rules of inference. *University of Chicago Law Review* 69, 1–133.
- Epstein, Lee and Andrew D. Martin (2014). *An Introduction to Empirical Legal Research*. Oxford University Press.
- Esienberg, Theodore , Talia Fisher, and Issi Rosen-Zvi (2011). Israel’s supreme court appellate jurisdiction : An empirical study. *Cornell Law Review* 96(4), 693–725.
- Teichman, Doron and Eyal Zamir (2014). Judicial decisionmaking: A behavioral perspective. In E. Zamir and D. Teichman (Eds.), *Oxford Handbook of Behavioral Economics and the Law*. Oxford University Press.
- Weinshall-Margel, Keren (2016). *Law and Ideology in Supreme Court Decision-Making: A Comparative and Quantitative Analysis*. Hebrew University Sacher Press.